

Original Created: V1.0 March 18, 2020
 This Version Created: V3.0 Sept 1st, 2020
 This Version Updated: V3.0 Sept 4th, 2020
 Nicola Barileto, Jan Oledan, Luca Zanotti (V2.4)
 Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

COVID Misinformation Codebook V3.0

This document describes the key fields and attributes of the ongoing COVID Misinformation tracking effort. Complete list of changes from V2.4 to V3.0 (current version) available on pg. 8 of this document.

Field	Description
ID	Simple numeric identifier.
Reported_On	URL of where the misinformation was reported (in a news article, blog, report). ***Please see the end of this table for instructions for misinformation that is not reported in articles.
Additional_Reporting	Include alternative URL if the article was reported in the AFP fact checking website .
Retrieved_from_1 (2, ..., n)	Include the URL of where the social media post is archived or mentioned. Most of the time, if the story reports a Twitter post, for example, they will mention the platform or include a screenshot of the Tweet or content. Limit one URL per variable. ***Please see the end of this table for instructions for misinformation that is not reported in articles.
Twitter_Reference	Captures whether the story mentions Tweets or Twitter. <ul style="list-style-type: none"> Equal to 1 if yes, 0 if no.
Direct_Post_1 (2, ..., n)	Include the direct URL of the social media post if it is still up and active. For example, direct link to a Tweet, Facebook post, or YouTube video. Limit one URL per variable. ***Please see the end of this table for instructions for misinformation that is not reported in articles.
Title	Main title of the news article, blog, report from the previous field.
Publication_Date	Date that the news article, blog, report was published.
Entry_Date	Date that the news article, blog, report was added to the sheet.

Original Created: V1.0 March 18, 2020
 This Version Created: V3.0 Sept 1st, 2020
 This Version Updated: V3.0 Sept 4th, 2020
 Nicola Barileto, Jan Oledan, Luca Zanotti (V2.4)
 Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

Primary_Country	<p>Where the disinformation is spread according to the article.</p> <ul style="list-style-type: none"> For Tweets or specific social media posts, it's the location the article references as being the targeted location or the coder's best guess as to what that is if none is referenced. If it is ambiguous, place "Ambiguous" as an option.
Secondary_Country	<p>Include all other countries or regions mentioned in the article (if any), separated by commas. E.g. Canada, United States</p>
Primary_Language	<p>What language is the piece of disinformation being spread? Limit one language for this field.</p> <ul style="list-style-type: none"> If a specific screenshot or social media post is available, use the language that is provided. If the article mentions the disinformation spread in a different language (but with no example, like in the previous case), use this language.
Secondary_Languages	<p>Include all other languages mentioned in the article (if any), separated by commas. E.g. Spanish, Italian.</p>
Main_Narrative	<p>The type of story the piece of misinformation is pushing. What's the content of the story? Note: limit to one narrative per entry.</p>
<ul style="list-style-type: none"> Options 	<p>False cures and preventative measures – the story reports unproven cures or procedures to take to fight COVID-19, including misinformation pertaining to vaccine development and trials.</p> <p>Nature of the virus – the story refers to specific characteristics about the virus, who it affects (age or ethnic groups) or how it can be spread (e.g., elderly people are more hit, young people are less susceptible, it cannot be spread on surfaces, etc.). Includes stories which:</p> <ul style="list-style-type: none"> claim the virus has a supernatural origin, e.g. "COVID is a punishment by God for our sins." predict the virus (e.g. claims the coronavirus occurs every 100 years or that a book predicted the pandemic); claims that the virus is not real; attempt to rename the virus (e.g. referring to the virus as the "Wuhan Pneumonia"); spread misinformation on the mortality rate of the virus; or

	<ul style="list-style-type: none"> - claim that the virus is not a threat or not as serious as credible public health sources suggest it is (e.g. stories discounting WHO or CDC reporting on fatality rates). <p>Origin of the virus – the story reports unverified origins of the virus (e.g., it came from consuming bat soup, or it came from eating meat) or false information about sources of infection. Excludes stories that suggest COVID-19 is a designed organism (e.g. bio-weapon, originated from a U.S. lab, was brought over to China by the U.S. Army).</p> <p>COVID-19 status of individuals and groups – the story speculates on patient zero identities, individual or group status to include infection or fatality rates within specific populations, and celebrity test as well as recovery results (e.g. unconfirmed reports regarding Patient Zero in Kenya, avoid certain areas because there’s a large Chinese population that have the virus, or Tom Hanks using COVID-19 as an excuse to stay in Australia because he’s part of a child sex ring associated with Weinstein). Includes misinformation on the outcome of individuals diagnosed with the virus (e.g. burial areas or funeral homes overwhelmed by the number of deaths, overfilled hospitals, misreported mortality rates in specific countries, etc.), as well as misinformation on the recovery status of individuals (e.g. large numbers of recovered patients are donating their plasma).</p> <p>Government responses – the story reports false information regarding government or political responses to fight the outbreak (e.g. US national guard locking down the US, or Italian biocontainment protocols). The responses can be emergency, short-term responses or anything related to reopening, stimulus, new rounds of lockdowns, a second wave of the virus, etc. This category should also include any praises or criticisms regarding government responses from the government side. Stories about a return to normal fall under this category if the government is a clear actor (e.g. China announces a return to normal, Taiwan reopens closed stores). A general “return to normal” with no clear order from the government falls under the “other” category.</p> <p>Non-Government responses – the story reports false information regarding individual, nonprofit, social, and business responses pertaining to COVID-19. Includes misinformation on positive or negative responses taken by non-government actors. For example, any misinformation about reopening by such entities, the breaking of</p>
--	--

	<p>lockdown compliance by non-government actors (e.g. breaking quarantine or social distancing rules, refusing to wear masks, or holding religious gatherings during lockdown), or misinformation about individual protests and backlashes, as well as fake praise for fabricated about non-governmental responses to the pandemic, or claims of the community helping out individuals (e.g. supermarkets providing assistance or restaurants offering free deliveries), and so on etc. fall into this category.</p> <p>False diagnostic procedures – the story reports unproven diagnostic procedures for the virus (e.g. holding your breath is a test for carrying COVID-19).</p> <p>Weaponization or design – the story describes COVID-19 as a designed organism or an intentional part of a country’s bio-weapons plot (e.g. Chinese, American, Russian plots), or an unintentional consequence of genetically-engineered organisms. Also, includes claims that the virus is being used as a weapon, even if the narrative admits that the virus is naturally occurring.</p> <p>Other – for stories which may not fall in any of the previous narrative definitions.</p>
Narrative_Description	<p>Adds another level of detail to describe the entry in the <i>Main_Narrative</i> field, according to the proposed guidelines above. Provide specifics on what the content is about with details from the story.</p> <ul style="list-style-type: none"> • For example, if the main narrative is “Weaponization or design”, this could be “US registered a patent in 2003 and used the virus as a bioweapon against China”. • Mention if the narrative frames the virus in religious terms and how.
ChainMsg_or_Email	<p>Captures whether the disinformation was spread through peer-to-peer channels. Equal to 1 if the disinformation was spread through chain message or phishing email, and 0 otherwise.</p>
Motive	<p>What is the disinformation trying to accomplish? Given the content and context, record subjective judgement about what the producer of the information is trying to achieve with promoting the disinformation.</p>

<ul style="list-style-type: none"> Options 	<p>Fear – disinformation with no discernible purpose beyond stoking fear among the public.</p> <p>Profit – stories which are linked to a financial motive, e.g. selling a cure or preventative measure for COVID-19.</p> <p>Politics – stories that target specific political actors or groups (e.g., President Trump, Italians in the US, Asian-Americans) to either:</p> <ul style="list-style-type: none"> Discredit and weaken them (e.g. Xenophobic remarks, smears); or Help their political standing (e.g. COVID-19 is a deep-state plot). <p>Undermine target country institutions – efforts to attack governments, international organizations, or specific ministries/bureaucracies (e.g. UN, EU, NATO, Italian healthcare system, etc.).</p> <p>Downplay Severity – stories which try to paint COVID-19 and the pandemic as normal or low-consequence (e.g. it’s just like the flu, hospitals are totally fine, it’s no more infectious than other diseases).</p> <p>False Hope – stories trying to instill a sense of hope (e.g. X celebrity is turning their house into a hospital to take in patients, there’s already a cure, a vaccine has already existed).</p> <p>Help – stories spread by the source with the goal of being helpful (e.g. false cures and preventative measures spread with good intent). Use this category only when the reporting and/or context clearly indicate the misinformation narrative was spread with good intentions.</p> <p>Other – content which does not fit into the above, including sarcastic commentary and individuals trying to sell fake stories for the sake of it (e.g. the Tom Hanks story from above, Nostradamus predicted this years ago, Author X predicted this in their book).</p>
Motive_Description	Provides greater detail to justify the entry from the Motive field. Write down specifically who the actors are and what they are doing given the coding and the context (one example of “Undermine target country institutions” is “Russian efforts to weaken the US administration and create chaos in the US by passing false information regarding the US response”).

Original Created: V1.0 March 18, 2020
 This Version Created: V3.0 Sept 1st, 2020
 This Version Updated: V3.0 Sept 4th, 2020
 Nicola Bariletto, Jan Oledan, Luca Zanotti (V2.4)
 Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

Source	How do we characterize the group that is pushing this misinformation?
<ul style="list-style-type: none"> Options: 	<p>Individual actor – individuals or groups (“common people”) who push misinformation through phishing scams, social media.</p> <p>State sponsors – states who are pushing misinformation efforts.</p> <p>Companies – known companies who exacerbate or house misinformation.</p> <p>Media – known news websites, TV networks that push misinformation.</p> <p>Political actor – known individuals or groups who support a certain politician or party pushing the misinformation.</p>
Source_Description	Who specifically is pushing this misinformation?
<ul style="list-style-type: none"> Options: 	<p>If key field <i>Source</i> contains:</p> <p>Individual actor</p> <ul style="list-style-type: none"> If we know the specific actor or group, name them. If we do not know the specific actor or group, place “General Public.” Typically, we use this for reports that contain chain messages or phishing emails because we are mostly unsure who started the chain message. <p>State sponsors</p> <ul style="list-style-type: none"> If we know the specific country or state, name them. <p>Companies</p> <ul style="list-style-type: none"> If we know the specific company, name them. <p>Media</p> <ul style="list-style-type: none"> If we know the specific news or TV network, name them. <p>Political actor:</p>

Original Created: V1.0 March 18, 2020
 This Version Created: V3.0 Sept 1st, 2020
 This Version Updated: V3.0 Sept 4th, 2020
 Nicola Barileto, Jan Oledan, Luca Zanotti (V2.4)
 Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

	<ul style="list-style-type: none"> ● If we know the specific government representative (House representative, Senator, Member of Parliament, Minister), name them. ● If we know only of groups or supporters for a particular government representative, name them as such “<Government representative name> supporter.”
Distrib_Channel	On which platforms (social media or not) is the misinformation being spread? Record the platform(s) that were mentioned in the articles.
▪ Options:	Facebook, Twitter, WhatsApp, Kakao Talk, TikTok, Youtube, Podcasts, 4Chan, Website (Personal, Blog) etc. (Add along as needed)
Key_Words	List of a few keywords related to the article and the misinformation.
Summary	Short sentence describing the article and the specific misinformation that is being spread.
Coder	For internal use, to keep track of entry additions.

***Instructions for misinformation that is not reported in articles:

This is used when the coder finds, for example, a post or Tweet that given knowledge and background, is deemed untrue but is otherwise not reported formally on a news platform or website. Please fill the fields accordingly:

- **Reported_On:** input “Coder Investigation”
- If this is found on a non-news platform website: place the URL into **Retrieved_From**
- If this is found on Twitter, Facebook, or other social media: place the URL into **Direct_Post**

Other Tabs in the Spreadsheet:

Resources – a list of crowdsourced COVID-related datasets, dedicated fact-checking websites, collections of situational reports, and other resources for researchers.

Aliases – a sheet recording different nicknames in different languages and countries for “COVID-19” or “Coronavirus”

AFP Links – a sheet collecting and tracking the narratives from the AFP fact-checking website, coded in the same way as our schema.

Changes from V2.4 to V3.0:

- Eliminated Emergency responses and split it into government and non-government responses, defined as follows:
 - **Government responses** – the story reports false information regarding government or political responses to fight the outbreak (e.g. US national guard locking down the US, or Italian biocontainment protocols). The responses can be emergency, short-term responses or anything related to reopening, stimulus, new rounds of lockdowns, a second wave of the virus, etc. This category should also include any praises or criticisms regarding government responses from the government side. Stories about a return to normal fall under this category if the government is a clear actor (e.g. China announces a return to normal, Taiwan reopens closed stores). A general “return to normal” with no clear order from the government falls under the “other” category.
 - **Non-Government responses** – the story reports false information regarding individual, nonprofit, social, and business responses pertaining to COVID-19. Includes misinformation on positive or negative responses taken by non-government actors. For example, any misinformation about reopening by such entities, the breaking of lockdown compliance by non-government actors (e.g. breaking quarantine or social distancing rules, refusing to wear masks, or holding religious gatherings during lockdown), or misinformation about individual protests and backlashes, as well as fake praise for fabricated about non-governmental responses to the pandemic, or claims of the community helping out individuals (e.g. supermarkets providing assistance or restaurants offering free deliveries), and so on etc. fall into this category.
- **Also expanded on some definitions of older categories of main narratives:**
 - **COVID-19 status of individuals:** Specified that this includes speculation as to the COVID status of defined groups, e.g. children are 100% immune.
 - **False cures and preventative measures:** Expanded the definition to also include misinformation pertaining to vaccine development and trials.
 - **Nature of the Virus:** Clarified the definition to also include claims that the virus is not real, attempts to name the virus, as well as misinformation that claims the virus is not a threat or not as serious as reliable public health agencies report
 - **Weaponization or design:** Expanded the definition to include claims of uses of the virus as a weapon, even if the same narrative admits that the virus is naturally occurring.

Changes from V2.3 to V2.4:

- Renamed “*Country*” to “*Primary_Country*” – limit to one option.

Original Created: V1.0 March 18, 2020
 This Version Created: V3.0 Sept 1st, 2020
 This Version Updated: V3.0 Sept 4th, 2020
 Nicola Barileto, Jan Oledan, Luca Zanotti (V2.4)
 Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

- Added “*Secondary_Country*” – include all other countries and regions mentioned in the article (if any), separated by commas.

Changes from V2.2 to V2.3:

- Changes to **Main_Narrative** field option “Nature of the virus” – includes stories which claim the virus has a supernatural origin, e.g. “COVID is punishment by God for our sins.”
- Changes to **Narrative_Description** field – mention if the narrative frames the virus in religious terms and how.
- Additions and refining definitions in the **Motive** field:
 - Downplay Severity: stories which try to paint COVID-19 and the pandemic as normal or low-consequence (e.g. it’s just like the flu, hospitals are totally fine, it’s no more infectious than other diseases)
 - False Hope: stories trying to instill a sense of hope (e.g. X celebrity is turning their house into a hospital to take in patients, there’s already a cure, a vaccine has already existed)
 - Help: stories spread by the source with the goal of being helpful (e.g. false cures and preventative measures spread with good intent). Use this category only when the reporting and/or context clearly indicate the misinformation narrative was spread with good intentions
 - Politics: emphasize targeting specific individuals or groups, change definition to include examples of xenophobic remarks and smears
 - Undermine target country institutions: refine definition to draw examples on larger entities (e.g. entire healthcare systems, administrations, multinational organizations such as the WHO, NATO, EU)
 - Other: disinformation that truly does not fall into the other options (e.g. Nostradamus predicted this years ago, Author X predicted this in their book)
- Refining definition of the **Motive_Description** field: justify the entry in the motive field. Why is the particular motive for the piece of disinformation coded that way?
- **Added process for disinformation that is not reported in articles** (i.e. coder finds a Tweet or post that given knowledge and background, is untrue)
 - Remember, we’re trying to code narratives and not individual content – so before adding, double check and make sure a particular narrative hasn’t been recorded.
 - **Reported_On**: Coder Investigation
 - If this is found on a website: place the URL into **Retrieved_From**
 - If this is found on Twitter, Facebook, or other social media: place the URL into **Direct_Post**
- **Added** description of other tabs in the spreadsheet.

Original Created: V1.0 March 18, 2020
This Version Created: V3.0 Sept 1st, 2020
This Version Updated: V3.0 Sept 4th, 2020
Nicola Bariletto, Jan Oledan, Luca Zanotti (V2.4)
Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

Changes from V2.1 to V2.2:

- Added “Other” as a *Main_Narrative* option – for stories which may not fit in the other main narrative options.
- Changes to “COVID-19 status of individuals” in the *Main_Narrative* field – include stories that describe individuals and groups of people. For example, “avoid certain areas because there’s a large Chinese population that have the virus.”
- Renamed “*Language*” to “*Primary_Language*” – limit to one option.
- Added “*Secondary_Languages*” – include all other languages mentioned in the article (if any), separated by commas.

Changes from V2.0 to V2.1:

- Include governments as a group under the Politics motive. Also include self-promotion as part of the Politics motive.
- Added *Country* and *Language* fields.
 - *Country*: Where the disinformation is spread according to the article.
 - *Language*: What language is the piece of disinformation being spread?

Changes from V1.1 to V2.0:

- Added *ID* field as a simple numeric identifier.
- Added *Additional_Reporting* field: Include alternative URL if the article was reported in the AFP [fact checking website](#).
- Changed *Date* field to *Publication_Date*: date that the news article, blog, report was published.
- Added *Entry_Date* field: date that the news article, blog, report was added to the sheet.
- Changed field name and options of *Narrative* field to *Main_Narrative* and added another category, *Narrative_Description*.
 - Options to *Main_Narrative* have been redefined.
 - *Narrative_Description*: adds additional detail to the *Main_Narrative* field.
- Added *ChainMsg_or_Email* field: binary field, equal to 1 if the disinformation was spread through chain message or phishing email, 0 otherwise.
- Added *Motive* and *Motive_Description* fields.
 - *Motive*: what is the disinformation trying to accomplish, given the content and context?
 - *Motive_Description*: adds additional detail to the *Motive* field.
- Changed field name of *Source_Subcat* to *Source_Description*.

Original Created: V1.0 March 18, 2020

This Version Created: V3.0 Sept 1st, 2020

This Version Updated: V3.0 Sept 4th, 2020

Nicola Bariletto, Jan Oledan, Luca Zanotti (V2.4)

Samikshya Siwakoti, Isra Thange, Anne Wen, Prof. Shapiro (V3.0)

- Added *Coder* field: for internal use, to keep track of entry additions.
- Added *Retrieved_from_1* (2, ..., n), *Direct_Post_1* (2, ..., n), *Twitter_Reference* fields.